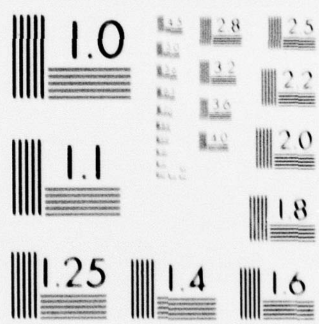
AD-A060 486 YALE UNIV NEW HAVEN CT DEPT OF PSYCHOLOGY F/G 5/10
A MODEL FOR PROCEDURAL LEARNING.(U)
1978 J R ANDERSON N00014-77-C-0242
UNCLASSIFIED NL

1 OF 1
AD
A060 486

END
DATE
FILMED
1-79
DDC



MICROCOPY RESOLUTION TEST CHART

(9) Final rept. 1 Mar 77-30 Sep 78,

Final Report

Period: Mar. 1, 1977 to Sept. 30, 1978
Contract No. (15) N00014-77-C-0242
NR No. 154-399

Project Title: (6) A Model for Procedural Learning.

Principle Investigator: John R. Anderson
Department of Psychology
Yale University

(10) John R. Anderson

This is the final report on the research performed at Yale University. The research will continue at Carnegie-Mellon University.

Summary of Progress to Date (12) 8p. (11) 1978

We have completed construction of the ACTF system described in the original proposal. There are four major mechanisms at operation that give this system the ability to generate and modify productions. One mechanism is referred as production designation. This mechanism is the means by which one production can designate the creation of another production. The basic technique is for the designation production to point to two structures in memory, one as the pattern for the condition of a designated production and one as the pattern for the action of the designated production. The memory structures that serve as patterns could either be created by a designating production or could already have existed in memory and simply have been retrieved by the designating productions. Much of the work on designation has concerned developing facilities that would enable this designation mechanism to operate flexibly in a variety of learning situations. Among these facilities are means for replacing nodes in the memory structure by global and local variables, and facilities for replacing pattern nodes by other nodes.

A second mechanism of learning is strengthening. Each production has a strength measure associated with it which is a measure of how successful it has been in the past. The strength of a production controls its probability if being selected and applied in a particular situation. Thus strength provides ACT with a means of selecting among competing procedures and of allowing past experience to vote as to what is the optimal procedure for a particular situation. The principle technical developments involving strength have been in coordinating it with the designation process and a specificity ordering on productions. We allow productions which have just been designated to be selected and applied independent of their strength. There is a "specificity ordering" on productions such that productions with more conditions to satisfy (before the action applies) take precedence in application over less specific productions. This specificity ordering is needed to enable ACT to encode exceptions to rules. We had to face the case of what to do when a weaker but more specific production was put into competition with a stronger but less specific production. The solution that seems to work involves having each production go through two passes of filtering. In the first pass, strength only operates. This is the probabilistic selection

phase in which each production is selected with a probability that depends on its strength. Productions which survive this selection hurdle are then chosen among according to specificity, independent of strength.

The third learning mechanism is generalization. This involves the comparison of two productions which perform identical actions. An attempt is made to generalize the conditions of the two productions in such a way that a new condition is created that will apply in all the situations of the original two porductions plus additional situations. A new production is created with this condition and the action of the original production. This generalization process creates a new condition from the original two by means of replacing nodes in which the two original productions differ by variables and by means of deleting structures on which they disagree. The technical problems tackled with respect to generalization concerned making it computationally efficient and restricting it to those circumstances where the original two productions are reasonably similar. There is no general solution to the efficiency problem but we have developed some heuristics that have cut CPU time by a factor of 10 in practical problems we have tried. To focus generalization on close matches simply requires specifying how much of a mismatch should be tolerated. We have been experimenting with a number of formulas but have no firm conclusions to report as yet.

The fourth learning mechanism is discrimination. This involves adding further constraints to a production's condition so as to limit the range of situations in which it can apply. This is the means by which overgeneral productions can be corrected. The principle technical problem with this mechanism has been in developing constraints on when discrimination should occur. The original proposal had intended to allow discrimination to occur for all productions that were active. However, this wound up producing a great many harmless but useless discriminant productions. This was very expensive in computation time and so inefficient that we concluded it was implausible psychologically. We now only apply the discrimination process to productions for which there is evidence of misapplication. Combined with the strength mechanism this appears to offer good performance.

We have been testing these mechanisms, individually and in combination, by means of a series of mini-simulations. The outcome of these tests and a much fuller description of the basic learning mechanisms is contained in Technical Report 77-1, A theory of the Acquisition of Cognitive Skills. The mini-simulations we have looked at so far have involved development of arithmetic skills, programming skills, and linguistic skills. They have supported our conjectures about the basic learning mechanisms.

We have developed a special version of the ACT system which is concerned with modelling efficiently the interaction between strengthening, generalization, and discrimination. This program is being successfully used to simulate the research on schema abstraction or prototype development.

During this period we have also completed development of all the parts of the ACTG system. In the ACTG it is possible to inspect much of one's

procedural knowledge. This proves to be important in modelling planning and other strategic uses of knowledge in problem solving. We have also completed the data flow matcher which allows data to select productions rather than productions to select data.

Our experimental work to date has mainly focused on testing our ideas about the behavior of the strengthening mechanism. This has involved looking at the effects of practice and attention on various mental procedures. The procedures we have chosen to look at have involved attribute judgments (e.g., dishonest) evaluating algebraic expressions, and making of syntactic judgments about symbolic expressions such as those found in the programming language LISP. Except for the algebraic task we have been able to find clear evidence for a speed-up of a procedure that generalizes across specific problems. The algebraic task showed no differential speed-up of ability to evaluate practiced expressions. We suspect this may reflect the overlearning of algebraic skills in our subjects.

We are currently looking at the interaction between strength and attention. La Berge has shown for perceptual tasks that amount of practice is not as important when one is attending to the task. We are trying to see if we can reproduce this finding, which is predicted by ACT, for the cognitive tasks we are concerned with.

One of the content areas that we had proposed to apply ACT to was the acquisition of the kind of textual information that is used in domains such as geography. The main idea was to teach ACT the kind of study skills that are required to be a good text learner. However, as we worked further on the problems it became clear that the process of acquiring a study skill was not where the significant learning issue lay. Rather it lay in what ACT's recommendations were about good study techniques for text material. These recommendations are based on past research with laboratory materials, but have never been put to the test with material as complicated as an actual textbook. The basic prediction of the ACT theory is that learning will be facilitated by any operation that enriches to-be-learned material with redundant details. ACT does not expect much of an effect of this elaboration on short-term retention but does expect an advantage for long-term retention, which usually is what is important.

We have set out to test this idea with a number of chapters from actual college texts. We have contrasted students' retention after reading summaries of the texts (about a quarter of the original length) with retention after reading the actual texts. Since the original texts differed from the summaries basically in redundant detail, this seemed like an excellent way to put our hypothesis to test. Subjects are given equal study time on both types of text. To our surprise, we found subjects immediately and a week later better on the summaries. We are just now gathering six month retention data which we regard as the true test of whether there is an effect on long-term retention. We have since gathered evidence that the advantage for the summary groups was due to difficulties subjects had in extracting or determining the main points of the text. The difference between the groups disappears when the main points are underlined in the

text. However, this outcome has really brought into focus for us the issue of what the value is of a traditional textbook.

The following Technical Reports were distributed during the project's residence at Yale:

77-1 A Theory of the Acquisition of Cognitive Skills

77-2 Design of a Production System for Cognitive Modelling

77-3 Status of Arguments Concerning Mental Imagery

78-1 Complex Learning Processes

Distribution List

Dr. M.J. Farr (5)
Dr. Charles E. Davis
Dr. Harold F. O'Neil, Jr.
Dr. William Montague
Dr. Worth Scanland
Dr. James Lester
Dr. John D. Ford, Jr.
Mr. Arnold I. Rubinstein
Dr. Norman J. Kerr

Technical Report 77-1: A Theory of the Acquisition fo Cognitive Skills
Contract No.: N00014-77-C-0242, NR No. 154-399

The paper describes the ACT theory of learning. The theory is embodied as a computer simulation program that makes predictions about human learning of various cognitive skills such as language fluency, study skills for social science texts, problem-solving skills in mathematics, and computer programming skills. The learning takes place within the ACT theory of the performance of such skills. This theory involves a propositional network representation of general factual knowledge and a production system representation of procedural knowledge. Skill learning mainly involves addition and modification of the productions. There are five mechanisms by which this takes place: designation, strengthening, generalization, discrimination, and composition. Each of these five learning mechanisms is discussed in detail and related to available data in procedural learning. Designation is the process by which one production can designate another. The power of the designating production is postulated to vary with sophistication of the learner in the domain to be learned. Strengthening is the process by which successful productions gradually acquite more control over the processing resources. This mechanism is related to the available data about how a skill gradually becomes automatic. Generalization is the process by which productions extend their range of application beyond the domain for which they were originally designated. This mechanism nicely accounts for phenomena in language acquisition and concept development. Discrimination is a coercive mechanism by which the range of overgeneral productions is restricted. This mechanism is used to explain phenomena in language acquisition and how problem-solving skills, such as in mathematics, evolve specialized submethods. Composition is the process by which multiple productions can combine into a single production. It explains the Einstellung effects in problem-solving. Finally, we discuss the difficulties we have encountered working with the current ACT learning systems and the projected changes in ACT to deal with these problems.

Technical Report 77-2: Design of a Production System for Cognitive Modelling #1
Contract No.: N00014-77-C-0242
NR No. 154-399

Abstract

It is conjectured that a good cognitive psychology theory will lead to a good artificial intelligence (AI) program. If this is true there should be a convergence of psychological and AI considerations in theory construction. This convergence is illustrated in terms of ACT, a computer simulation model of cognitive processes. Separate AI and psychological considerations are used to motivate the decision to design ACT as a production system operating on an associative network data base. Similar motivation is provided for other features of ACT implemented within this framework. These features include the use of a propositional structure for the associative network, a spreading activation process operating on the network, the simulated ability to execute several procedures in parallel, and the use of strength measures to select among competing productions and competing paths in the network.

Technical Report 77-3: The Status of Arguments Concerning Representations for Mental Imagery
Contract No.: N00014-77-C-0242
NR No. 154-399

Abstract

A review is provided of the recent debates over whether pictorial-like or propositional-like representations are most appropriate for visual imagery. The argument for a propositional representation has largely taken the form of an attack on the logical coherence of pictorial representations. These attacks have not been valid; one can develop a coherent dual-code model involving pictorial and verbal (non-propositional) representations. On the other hand, empirical demonstrations that are claimed to support pictorial representations fail to provide evidence that would discriminate such representations from propositional ones. It is argued that the failure of the anti-pictorial and the pro-pictorial arguments stems from a fundamental indeterminancy in deciding issues of representations. It is shown that wide classes of different representations can be made to yield identical behavioral predictions. In particular, this potential for mutual mimicry holds between propositional and dual-code (pictorial-plus-verbal) models. If one considers criteria such as parsimony and efficiency in addition to prediction of behavior, it may be possible to establish further constraints on representation. In particular, it may be possible to establish whether there are two codes, one for visual information and one for verbal, or whether there is a single abstract code. However, the conclusion of this paper is that, barring decisive physiological data, it will not be possible to establish the character of an internal representation--e.g., whether it is pictorial or propositional.

Abstract

This paper describes the ACT theory of the learning of procedures. ACT is a computer simulation program that uses a propositional network to represent knowledge of general facts and a set of productions (condition⇒action rules) to represent knowledge of procedures. There are currently four different mechanisms by which ACT can make additions and modifications to its set of productions as required for procedural learning: designation, strengthening, generalization, and discrimination. Designation refers to the ability of productions to call for the creation of new productions. Strengthening a production may have important consequences for performance, since a production's strength determines the amount of system resources that will be allocated to its processing. Finally, generalization and discrimination refer to complementary processes that produce better performance by either extending or restricting the range of situations in which a production will apply. Each of these four mechanisms is discussed in detail and related to the available psychological data on procedural learning. The small-scale simulations of learning provided as examples are drawn from the domains of language processing and computer programming, since our ultimate goal is for ACT to learn the complex procedures required in such domains.